References

1 Parmley, S. F. and Smith, G. P. (1988) *Gene* 73, 305–318.
2 Ladner, R. C. and Guterman, S. K. (1990) *Patent Appl.* WO90/02809
3 Ladner, R. C., Guterman, S. K., Roberts, R. L., Markland, W., Ley, A. C. and Kent, R. B. (1993) *US Patent No.* 5223409
4 Clackson, T. and Wells, J. A. (1994) *Trends Biotechnol.* 12, 173–183
5 Lane, D. P. and Stephan, C. W. (1993) *Curr. Opin. Immunol.* 5, 268–271
6 Cunningham, B. C., Lowe, D. G., Li, B., Bennett, B. D. and Wells, J. A. (1994) *EMBO J.* 13, 2508–2515
7 Ladner, R. C. in *Phage Display of Peptides and Protein Domains: A Laboratory Manual* (Kay, B., Winter, J. and McCafferty, J., eds), Academic Press (in press)
8 Creighton, T. E. and Charles, I. G. (1987) *Cold Spring Harbor Symp. Quant. Biol.* 52, 511–519
9 Brendt, K. D., Guntert, P. and Wuthrich, K. (1993) *J. Mol. Biol.* 234, 735–750
10 Hynes, T. R., Randal, M., Kennedy, L. A., Eigenbrot, C. and Kossiakoff, A. A. (1990) *Biochemistry* 29, 10018–10022
11 de Vos, A. M. *et al.* (1992) *Biochemistry* 31, 270–279
12 Balass, M., Heldman, Y., Cabilly, S., Givol, D., Katchalski-Katzir, E. and Fuchs, S. (1993) *Proc. Natl Acad. Sci. USA* 90, 10638–10642
13 Goodson, R. J., Doyle, M. V., Kaufman, S. E. and Rosenberg, S. (1994) *Proc. Natl Acad. Sci. USA* 91, 7129–7133
14 Cwirla, S. E., Peters, E. A., Barrett, R. W. and Dower, W. J. (1990) *Proc. Natl Acad. Sci. USA* 87, 6378–6382
15 Pinilla, C., Appel, J. A. and Houghten, R. A. (1993) *Gene* 128, 71–76
16 McLafferty, M. A., Kent, R. B., Ladner, R. C. and Markland, W. (1993) *Gene* 128, 29–36
17 Hoess, R. H., Mack, A. J., Walton, H. and Reilly, T. M. (1994) *J. Immunol.* 153, 724–729
18 Leatherbarrow, R. J. and Salcinski, H. I. (1991) *Biochemistry* 30, 10717–10721
19 Koivunen, E., Wang, B. and Ruoslahti, E. (1994) *J. Cell Biol.* 124, 373–380
20 Koivunen, E., Wang, B. and Ruoslahti, E. (1995) *Bio/Technology* 13, 265–270
21 Dennis, M. S. and Lazarus, R. A. (1994) *J. Biol. Chem.* 269, 22129–22136
22 Dennis, M. S. and Lazarus, R. A. (1994) *J. Biol. Chem.* 269, 22137–22144
23 Clackson, T. and Wells, J. A. (1995) *Science* 267, 383–386
24 Feng, S., Chen, J. K., Yu, H., Simon, J. A. and Schreiber, S. L. (1994) *Science* 266, 1241–1247
25 Scott, J. K. and Smith, G. P. (1990) *Science* 249, 386–390
26 Devlin, J. J., Panganiban, L. C. and Devlin, P. E. (1990) *Science* 249, 404–406
27 Kay, B. K., Adey, N. B., He, Y-S., Manfredi, J. P., Mataragnon, A. H. and Fowlkes, D. M. (1993) *Gene* 128, 59–65
28 Roberts, B. L. *et al.* (1992) *Proc. Natl Acad. Sci. USA* 89, 2429–2433
29 Smith, G. P., Schultz, D. A. and Ladbury, J. E. (1993) *Gene* 128, 37–42
30 Folgori, A. *et al.* (1994) *EMBO J.* 13, 2236–2243
31 Meola, A. *et al.* (1995) *J. Immunol.* 154, 3162–3172

Progress in anti-HIV structure-based drug design

Michael J. Gait and Jonathan Karn

The course of drug development for the treatment of HIV-1 infection and AIDS is being revolutionized by high-resolution structures of essential viral proteins. We survey the impact on drug design of the recently elucidated structural knowledge of two essential enzymes, reverse transcriptase and protease, and three new targets, the viral integrase and the gene regulatory protein–RNA interactions, Tat–TAR and Rev–RRE.

The speed of drug development for the treatment of human immunodeficiency virus (HIV) infection is unprecedented. In the 12 years since the discovery of the virus, four antiviral drugs have been licensed – three of these within the past two years – and scores more are undergoing clinical trials. However, despite this enormous effort, effective regimens for the prevention of the progression to clinical symptoms of acquired immune deficiency syndrome (AIDS) and the treatment of AIDS patients have not yet been developed. Typically, antiviral drugs produce a marked decrease in circulating HIV and an increase in $CD4^+$ T lymphocytes as a result of their anti-HIV properties. Unfortunately, these beneficial effects are short-lived, owing to the rapid emergence of drug-resistant strains.

Current strategies for 'trapping' the virus are based on finding combinations of drugs that are more effective than monotherapy. In theory, combinations of drugs aimed at multiple targets could reduce virus replication sufficiently to make the emergence of resistant variants highly unlikely. Alternatively, it may

M. J. Gait and J. Karn are at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge, UK CB2 2QH.

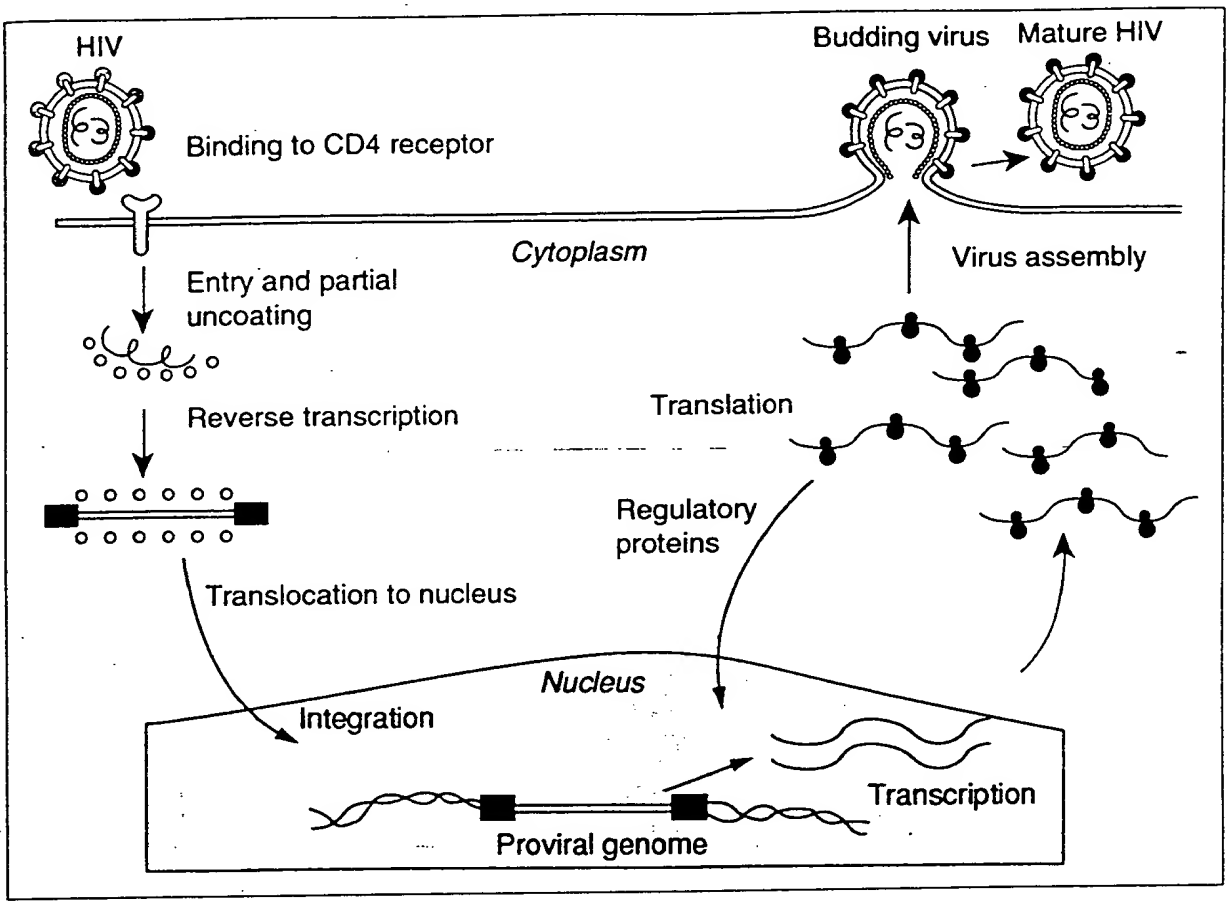


Figure 1
Replication cycle of HIV.

be possible to find combinations of drugs that, although directed to the same molecular targets, show a sufficiently different spectrum of escape mutations to prevent cross-resistance. For either strategy to be effective, it is essential to develop new and increasingly effective drugs.

The majority of clinical trials have involved drugs that have been found by traditional screening methods, and that are directed against either the HIV-1 reverse transcriptase (RT), or the HIV-1 protease. However, the recent availability of high-resolution three-dimensional structures for both these targets is rapidly changing the course of drug discovery. Although *de novo* drug design is still in its infancy, structural information is playing an important role in improving the properties of lead compounds, especially in the development of drugs directed against the HIV-1 protease. In addition, recent progress in the structural elucidation of further HIV targets, such as the viral integrase and regulatory proteins, is providing a basis for the development of potential novel clinical drugs.

Virus replication cycle

HIV and related lentiviruses have replication cycles that are typical of all retroviruses. The major biochemical steps in virus growth are shown in Fig. 1. It is convenient to think of viral replication as comprising four distinct stages: infection; reverse transcription and integration; viral-gene expression; and virus assembly and maturation. Soon after it has infected the host cells, the viral RNA genome is reverse-transcribed into a DNA copy that is transported to the nucleus and integrated at random sites in the host chromosome. Once integrated, the proviral genome is subject to transcriptional regulation by the host cell, as well as by its own transcriptional control mechanisms. The later stages of the life cycle involve the expression of the viral genes, and the eventual assembly and release of virus particles.

In principle, any essential step in the virus replication cycle presents a suitable target for drug discovery. In practice, drug discovery programmes have targeted a few of the HIV enzymes (RT, protease, integrase) and the regulatory proteins (Tat, Rev, Nef), rather than virion structural and accessory proteins (Env, Gag, Vif, Vpr, Vpu), in order to take advantage of convenient *in vitro* assays, as well as the vast experience of pharmaceutical researchers in the design of enzyme inhibitors.

Reverse transcriptase

RT enters the host cell with the RNA and the other virion proteins. The enzyme carries out several reactions, including reverse transcription of the RNA into cDNA, conversion of the newly synthesized DNA strand into duplex DNA, and removal of the unused RNA by an intrinsic RNaseH activity. Because RT—

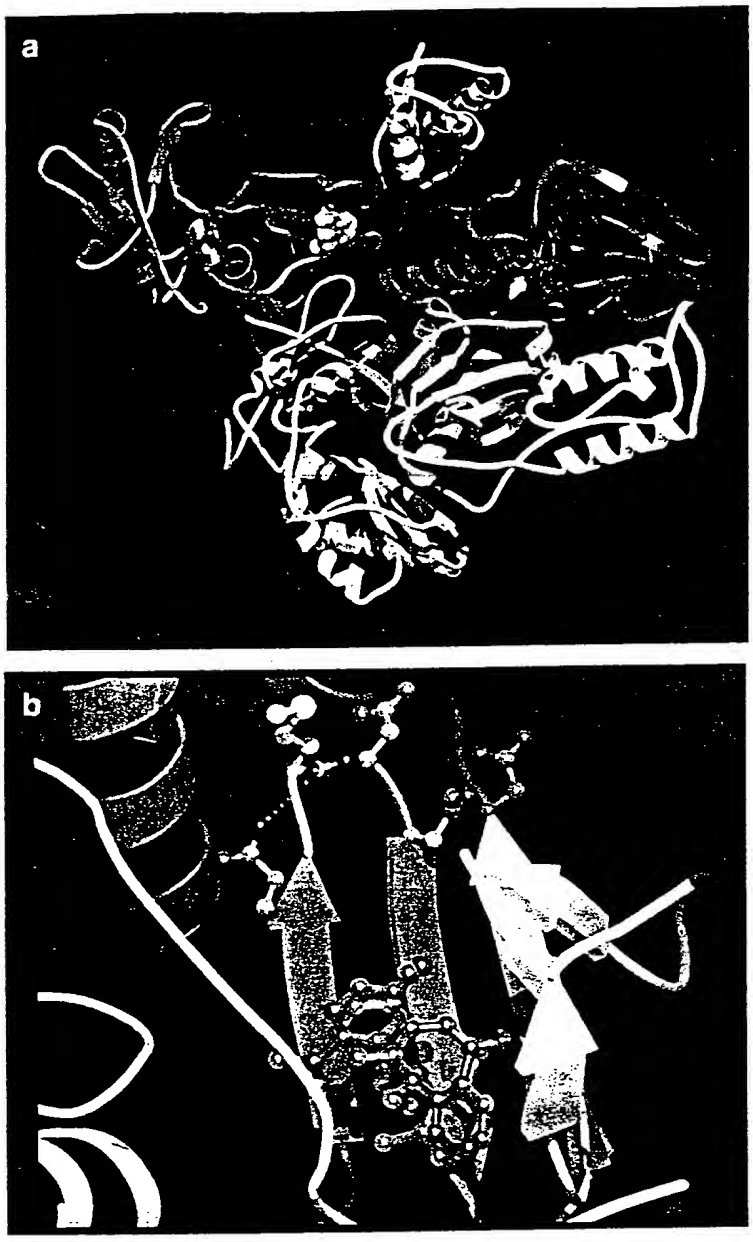


Figure 2

(a) Ribbon diagram of the reverse transcriptase (RT) p66/p51 heterodimer bound by nevirapine (space-filled orange)[5]. Colours represent domains of the p66 subunit (saturated) and p51 subunit (pastel) defined as 'fingers' (yellow), 'palm' (green), 'thumb' (pale yellow), 'connection' (red) and RNase H (purple). (b) Superposition of nevirapine (red) and α-APA (green) in the non-nucleoside inhibitor (NNI) binding site of RT. The catalytic residues Asp110, Asp185 and Asp186 are shown as atom-coloured ball-and-stick representations (top right) with Gln182 and Met184 (top left).

lacks an error-correction mechanism, the enzyme shows poor fidelity compared to cellular DNA polymerases, and introduces between five and ten mismatches per genome per round of replication. The high error-rate of reverse transcription helps to explain the wide variations in sequence among HIV strains, even within the same infected individual, as well as the rapid selection of drug-resistant mutants during treatment.

HIV-RT is processed from the polyprotein gene product of the *gag–pol* gene complex to give a p66–p51 heterodimer. The two subunits differ not only in their configurations, but also because p66 carries the catalytic domain for RNaseH in addition to the polymerase domain. It has taken many years of effort to obtain the crystal structure of RT, either on its own, or in complexes with drugs or model DNA substrates[1–5]. All of the structures have shown that the shape of RT is reminiscent of a right hand (Fig. 2a) with DNA threaded between the 'fingers' and a 'thumb' formed by the p66 subunit. The 'palm' subdomain of the p66 subunit contains the polymerase active site and the binding site for a number of non-nucleoside inhibitors (NNIs).

Practically all of the known RT inhibitors were found by screening before any structural information on RT was available. The four clinically approved nucleoside analogues [3′-azido-3′-deoxythymidine (AZT); 2′,3′-dideoxycytidine (DDC); 2′,3′-dideoxyinosine (DDI); and (–)-β-L-2′,3′-dideoxy-3′-thiacytidine (3TC)] act as chain terminators during DNA synthesis, in a similar manner to the chain terminators that are commonly used in DNA sequence analysis. Each of these modified nucleosides is converted into its corresponding triphosphate by intracellular kinases, and is subsequently incorporated into DNA by RT. Since the nucleoside analogues lack a 3′-OH group that can be phosphorylated, incorporation of the next nucleotide is blocked. High-resolution structures of RT that has been co-crystallized with nucleoside analogues have not yet been obtained, but the modes of interaction of these analogues with the RT complex can be inferred from docking simulations of a 2′-deoxynucleoside 5′-triphosphate (dNTP) into the polymerase active site[2,6].

It is already clear that the mutations in RT that confer drug resistance to the nucleoside analogues fall into two categories: some mutations are found close enough to the dNTP-binding site to suggest direct interference with drug binding; however, many are located on the surface of the DNA-binding cleft, and it is thought that these mutations lead to less-favourable incorporation of the drug compared with dNTPs because of altered primer-template positioning or conformational changes in the active site[6]. Clearly, a better understanding of the interactions of nucleoside analogues with RT will have an appreciable impact on further nucleoside-analogue drug design.

In addition to the nucleoside inhibitors, diverse classes of NNIs have been identified through drug-screening programmes. The crystal structures of complexes of four of these inhibitors (TIBO, HEPT, nevirapine and α-APA) with RT have revealed a common binding mode[5]. Each of these inhibitors binds to a hydrophobic pocket located in the 'palm' subdomain, some 10 Å from the polymerase active site (Fig. 2b). Surprisingly, the pocket is not present in the unliganded complex. Much of the energy of the binding of those compounds with RT comes from ring-stacking interactions with the residues Tyr181 and Tyr188. Inhibitor binding leads to a withdrawal of Trp229 from

the binding site and the flipping of the two Tyr side chains to allow entry of the NNI. This conformational change induces a repositioning of a three-stranded β-sheet in p66 containing the catalytic residues Asp110, Asp185 and Asp186 (Refs 3,4). Thus, the NNI-bound RT appears to be locked into an inactive conformation[7].

One important observation emerging from the structures of the four NNI complexes with RT is that the hydrophobic pocket is sufficiently flexible to permit the positioning of each inhibitor to be adjusted slightly, so that unfavourable contacts are relieved without altering the overall binding energy[5]. These subtle changes explain the differing patterns of resistance mutations seen with various classes of NNI (Ref. 8). This allows some scope for the use of structural information in designing modified NNIs that might induce more-slowly emerging RT resistance. Furthermore, since nucleoside triphosphates bind tightly to RT complexed with NNIs, and the two binding sites are within 10 Å of each other, another approach to drug discovery could be to design inhibitors thatcontain both types of functionality[7].

HIV protease

The second major target for HIV chemotherapy is the HIV aspartic proteinase (HIV protease). This essential enzyme cleaves the polyprotein produced from the *gag–pol* genes. Inhibition of the protease leads to defects in the virion-assembly pathway, resulting in the production of noninfectious virus particles.

HIV protease is structurally and mechanistically related to mammmalian and other microbial aspartic proteases, such as pepsin, cathepsin D, renin and endothiapepsin. However, in contrast to cellular proteases, which are monomeric, the X-ray crystallographic structure of HIV-1 protease[9] has revealed it to be a homodimer (Fig. 3a). The active site is an elongated cavity formed between the two symmetric subunits. When the natural substrate or substrate analogues bind, two flexible flaps close around the active site. This brings two catalytic Asp residues close to the scissile amide linkage of the natural substrate. A crucial water molecule is trapped on the other side, between the flaps and the substrate, locking the enzyme into a 'closed' conformation. The first structure of a protease–inhibitor complex was published in 1989 (Ref. 10) and, by 1993, at least 160 structures of HIV protease–inhibitor complexes had been determined[11]. These structures have had considerable influence on the course of drug discovery and, thus, HIV protease represents the 'archetype' for structure-based drug design[11,12].

The first approaches to inhibitor design were based on peptide analogues of the natural substrate, and they were initiated well before the structure of HIV-1 protease was known. Of particular importance was the development of 'transition-state analogues', in which the scissile amide bond is replaced by a non-hydrolysable isostere. In the first protease inhibitor to reach clinical trials for the treatment of AIDS

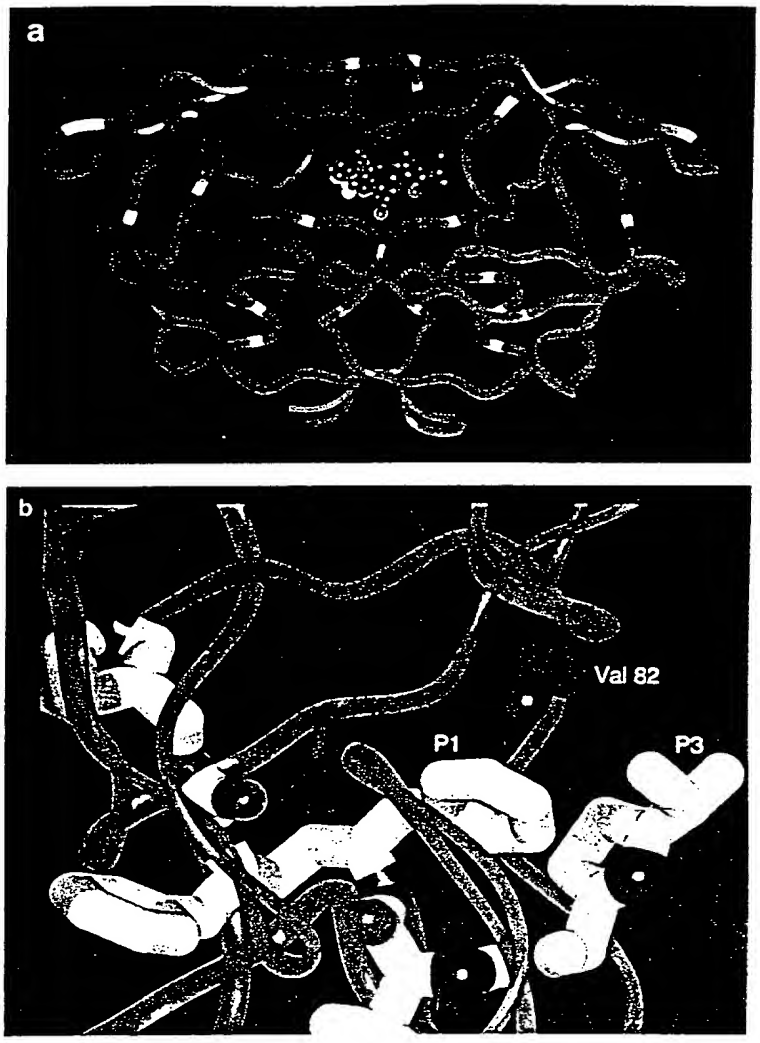


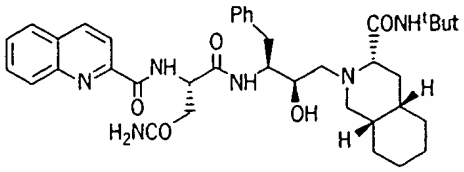
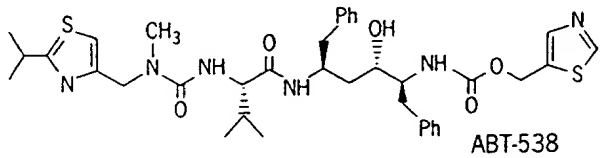
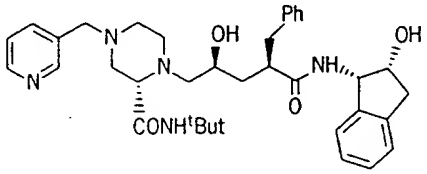
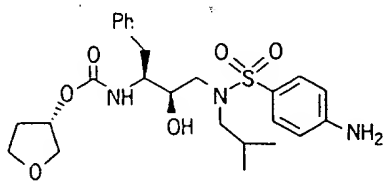
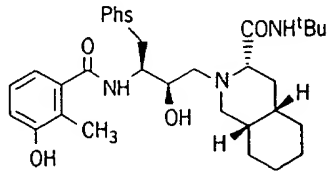
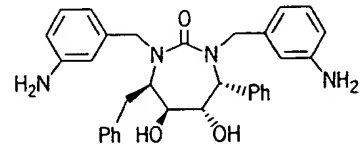
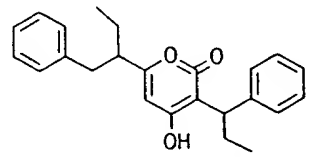
Figure 3

(a) Ribbon diagram of the homodimer of HIV protease with ABT-538 bound to the active site[15]. (b) Close-up of the additional hydrophobic contact designed between Val82 of HIV protease and the isopropyl substituent on the P3 thiazolyl group of ABT-538. Val82 is an important site for drug resistance[46]. The P1 benzyl group is also marked.

patients (Ro31-8959 from Roche[13]), hydroxyethylamine (-CH[OH]CH_2N-) replaces the peptide linkage (Table 1). Similarly, there is a hydroxyethylene (-CH[OH]CH_2-) isostere in L735524 (now known as MK-639), a clinical candidate from Merck[14]; this is also an example of the trend towards mimetics of shorter peptide substrates (four residues in this case) rather than the seven residues of the original peptide substrate.

A second class of protease inhibitors takes advantage of the C_2 symmetry observed in the crystal structure of the apo-enzyme. Compounds, such as A77003, from the Abbott drug-discovery programme, contain a C_2-symmetry axis together with a dihydroxyethylene transition-state isostere, and were founded firmly on computer-based modelling. One advantage of drug symmetry is that the synthetic pathway for manufacture is generally easier than for the synthesis of non-symmetrical drugs, such as Ro31-8959. Further development by Abbott has led to the development of a

Table 1. Structures of prominent HIV-protease inhibitors and their status in clinical trials

Structure		Developer	Status in clinical trials	Ref.
	Ro31-8959	Roche	Phase III	13
	ABT-538	Abbott	Phase II/III	15
	MK-639 (L735524)	Merck	Phase II	14
	VX-478	Vertex-Wellcome	Phase I	17
	AG1343	Agouron	Phase I	16
	XM412 (DMP 450)	Dupont-Merck	Phase I (terminated)	23
	U96988	Upjohn	Phase I	20

compound ABT-538 (Table 1 and Fig. 3a) that has much higher bioavailability in humans[15]. An important feature of ABT-538 is the inclusion of an additional N-terminal isopropyl group that is predicted to interact hydrophobically with Pro81 and Val82; this was confirmed by the crystal structure of the protease–ABT-538 complex (Fig. 3b).

Recently, there has been much effort to design smaller protease inhibitors (MW <600) that are easy to synthesize and have reduced clearance rates *in vivo*. The design of one such inhibitor, AG1343 from Agouron, was based on over 50 high-resolution crystal structures of protease–inhibitor complexes, and used powerful modelling programmes[16]. Both this analogue and VX-478, a potent and orally bioavailable inhibitor that emerged from a structure-based design programme at Vertex[17] (later licensed to Wellcome), incorporate a hydroxyethyl transition-state isostere and are modelled into only the four central amino acids of the substrate. A number of 4-hydroxycoumarin

derivatives, in which there are no peptide bonds, were developed independently by Parke-Davis[18,19] and Upjohn[20]. In each case, the leads were found by screening, but molecular modelling and X-ray structural data aided the development of improved compounds, such as the clinical candidate U96988 (Ref. 20) and the cycloalkyl ring derivatives of 4-hydroxycoumarins[21].

An important prediction from structure-based modelling was that replacement of the essential aspartate-based water molecule with a functional group would result in considerable entropy gain and, hence, improved binding. This concept was exploited in a series of compounds designed by Dupont-Merck, which were based on a readily prepared C_2-symmetrical cyclic urea structure in which the carbonyl of the urea replaces the water molecule that bridges the inhibitor and the closing protease flap[22]. The clinical lead is XM412 (Ref. 23), but this has been recently withdrawn from Phase I trials.

Early attempts at *de novo* design were largely based on the use of the 'DOCK' computer program[24]. With the development of better crystallographic data and improved energy minimization programs, it has been possible to predict the *in vitro* enzyme-inhibition potency of a series of protease inhibitors prior to their synthesis by calculating their 'interaction energy'. An example of this approach is the design of an inhibitor that was a precursor to the candidate MK-639 (Ref. 25), which is currently undergoing phase II clinical trials.

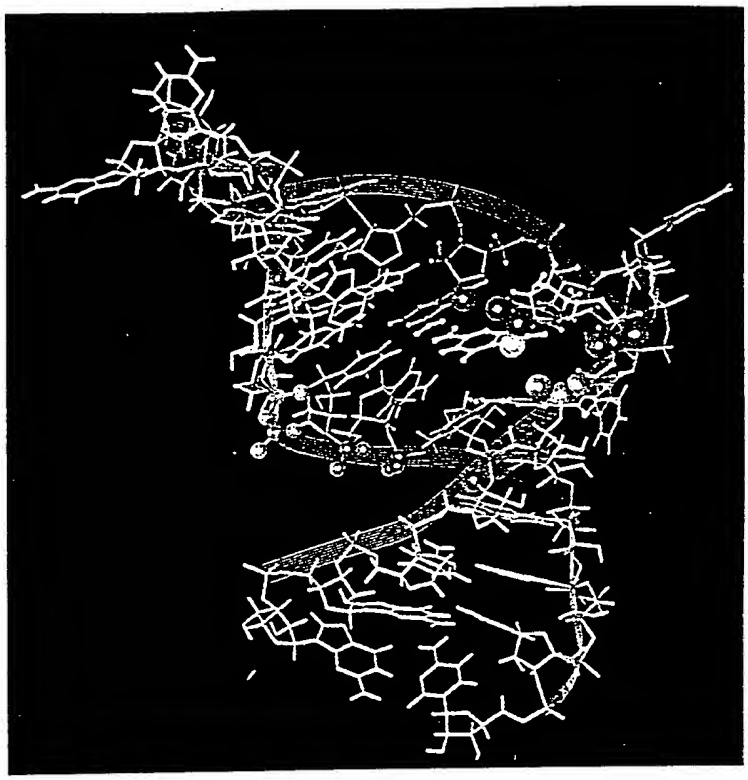


Figure 4

Model of the TAR region of HIV-RNA based on the NMR structure of TAR bound by a Tat peptide[36]. Residues U23 in the bulge (green) and G26 (yellow) have been shown by mutational and chemical studies to be critical for Tat binding[29], whilst two important A residues (A22 and A27) are shown in purple. Spheres delineate functional groups that may be contact sites with Tat.

Integrase

The integration of HIV-DNA into the host-cell genome is catalysed by HIV integrase in a two-stage reaction. In the first step, known as 3′-processing, two nucleotides (a G-T dinucleotide) are excised from the blunt-ended DNA duplex that is the product of reverse transcription. The reaction is a one-step *trans*-esterification in which integrase renders the cleaved phosphodiester bond susceptible to nucleophilic attack by the 3′-OH group of the dinucleotide (to form a cyclic dimer) or of water (to form a linear dimer). The second step is a strand-transfer reaction, in which the newly formed 3′-OH group of the recessed nucleotide derived from the first step attacks a phosphodiester bond in the target DNA, resulting in the formation of a branched structure. This second step has been shown to be reversible *in vitro*. The full-length 34 kDa protein is required for the 3′-processing, whereas strand transfer requires only the central core region. Nevertheless, there appears to be only one active site, which is located in the central core.

The development of *in vitro* assays that involve the use of synthetic oligonucleotide substrates has led to several drug-screening programmes (for example, at Merck)[26]. Although clinical candidate drugs have not yet emerged, a number of integrase inhibitors have been reported, including well-known intercalators such as mitoxantrone and ellipticines, and unrelated compounds such as caffeic acid phenylethyl ester[27].

A recent major advance has been the determination of the crystal structure of the catalytic domain (residues 50–212) of HIV-1 integrase at a resolution of 2.5 Å (Ref. 28). Crystallization of integrase had proved extremely difficult because of its insolubility and its tendency to aggregate. After a long search, a mutant protein that retained catalytic activity, but that had better solubility, was identified. The crystal structure revealed that a surface Phe had been replaced by Lys.

The integrase structure contains a central five-stranded β-sheet surrounded on three sides by numerous α-helices. Two of the three essential carboxylates (Asp64 and Asp116) are located in the β-sheet region. Another striking feature is the structural similarity of HIV-1 integrase to a family of polynucleotidyl transferases that includes ribonuclease H and the Holliday junction resolvase, RuvC. Modelling suggests that two or possibly four monomers probably combine in such a way as to bring together a pair of active sites with a spacing of about 15 Å (Ref. 28). Additional clues to the integrase mechanism will undoubtedly emerge once the structures of complexes between integrase and DNA are determined. Although structural studies are still at an early stage, it is clear that it should soon be possible to model compounds that have been obtained from screening programmes into the enzyme active site.

Regulatory proteins

The switch from latency to active viral replication in an infected cell is controlled by two viral proteins, the *trans*-activator protein (Tat) and the regulator of virion expression (Rev). These regulatory proteins play complementary roles in the HIV life cycle: Tat stimulates transcriptional elongation from the viral long terminal repeat (LTR); Rev is required for the efficient cytoplasmic expression of the mRNAs encoding the structural proteins of the virus.

In contrast to previously studied retroviral regulatory proteins that act through DNA elements, Tat and Rev are RNA-binding proteins that exert their effects through specific *cis*-acting viral-RNA regulatory sequences. Tat activity requires the *trans*-activation responsive region (TAR); in HIV-1, this is a 59-residue stem-loop RNA located at the 5′-end of all viral transcripts. TAR acts as a binding site for Tat, with recognition being centred around a U-rich bulge[29]. The apical loop region is thought to participate in *trans*-activation by acting as a binding site for essential cellular factors.

Rev activity is mediated through interaction with a region of viral RNA contained within the *env* gene which is known as the Rev-response element (RRE; Ref. 29). The RRE covers 351 residues, and is predicted to fold to give a number of RNA stem-loops as well as an unusually long central stem of about 90 base pairs. Rev binds initially to a high-affinity site that is characterized by a purine-rich 'bubble'; this contains bulged GG and GUA residues that are stabilized by non-Watson–Crick base pairs. At higher concentrations, Rev multimerizes on the RRE, packing along the length of the central stem[30].

The interactions between Tat and TAR, and between Rev and the RRE are critical for virus replication. Mutations that alter either RNA-binding site result in defective viruses. Furthermore, virus replication is strongly inhibited by the overexpression of TAR-RNA and RRE-RNA 'decoy' sequences that act as competitive inhibitors of regulatory-protein binding[31,32]. Therefore, it seems very likely that small molecules that inhibit the binding of Tat or Rev to their recognition sites will also have antiviral activity.

The search for drugs directed against these regulatory proteins is currently based on a variety of cellular and *in vitro* screens. An inhibitor of HIV-1 transcription was discovered by Roche following a screening programme using a cellular assay[33]. This benzodiazepine derivative was developed into a clinical candidate (Ro24-7429), but this was withdrawn from clinical trials as a result of toxicity. Surprisingly, the drug is not an inhibitor of the interaction of Tat with TAR-RNA, but appears to inhibit an unidentified cellular factor that is involved with the *trans*-activation pathway[34]. A second class of *trans*-activation inhibitor (3-keto-4,5-epoxy steroids) has also been identified using a similar cell-based screen[35], but these compounds also do not interfere directly with Tat action.

Detailed structural information about the RNA-binding sites for the regulatory proteins is now becoming available. We have recently obtained a high-resolution NMR structure of the Tat-binding site on TAR (Fig. 4; Ref. 36). The NMR studies demonstrate that the accessibility of critical functional groups recognized by Tat is enhanced by a local conformational rearrangement. This folding process can be induced by any ligands carrying a guanidinium group, including short basic peptides that encompass the basic region of Tat, and even the single amino acid analogue argininamide[37]. In the presence of these ligands, the stacking of the bulged residues U23 on A22 and C24 on U23 is disrupted, and A22 becomes juxtaposed to G26. This creates a binding pocket in which the guanidinium and ϵ-NH groups of the arginine are placed within hydrogen-bonding distance of the base functional groups G26-N^7 and U23-O^4 or U23-O^2, respectively.

Two possible avenues to drug discovery are suggested by these studies. First, the observation that the folded TAR structure provides a binding pocket for low molecular weight compounds such as argininamide, and the identification of a subset of the intermolecular contacts between TAR and the ligand[36,37], can be used to design functionality around the guanidinium group in order to increase the affinity and specificity of this interaction. Alternatively, compounds that 'lock' the flexible TAR structure into a conformation that is different from the one observed in the presence of the Tat-derived peptides could also be very effective inhibitors of the Tat–TAR interaction.

Work on the structure of the Rev–RRE complex is less advanced than that on the Tat–TAR complex. NMR analysis of the high-affinity binding site on RRE has confirmed that the non-Watson–Crick base pairs that are critical for Rev recognition are stabilized by the binding of peptide ligands[38,39]. The base pairs, together with a single looped-out U residue, form a binding pocket for Rev in the RNA major groove, and present an obvious target for inhibitor binding. Several inhibitors [e.g. the intercalator pyronin Y (Ref. 40) and the antibiotic neomycin B (Ref. 41)] of the Rev–RRE interaction have been found from screens based on RNA-binding assays. In addition, compounds that inhibit the Rev response in cellular screening assays and that have antiviral activity are being discovered; these include the 8-alkyl-2-(4-pyridyl)pyrido[2,3-d]pyrimidin-5(8H) compounds[42]. As NMR and crystallographic data on the Rev–RRE complex emerge, it will be possible to develop these leads further.

Drug resistance

The recognition that HIV mounts a chronic infection with a remarkably high level of virus replication even during the asymptomatic phases of the disease[14,43,44], has helped to explain why attempts to find effective drug monotherapy have been frustrated by the swift selection of drug-resistant viruses. When patients are treated with any of the currently available drugs as monotherapy, a reduction in viral load of about 10^2 is frequently observed within a two-week

period; however, these levels are rarely maintained over succeeding months, owing to the emergence of resistant virus strains.

The rate at which drug-resistant mutants emerge following therapy is a result of a variety of factors, including the efficacy of the drug itself, the viral replication and recombination rates, and the adverse effects of the resistance mutation on the viral growth rate. Most resistance mutants that are detected in clinical trials pre-exist in HIV infected individuals[43] but, because these mutations reduce virus growth rates, they will only emerge as dominant strains during drug treatment. It has been difficult to predict the patterns of emerging drug resistance in advance of clinical studies. In the case of RT and protease, many of the resistance mutants are found at sites that are far removed from the drug-binding site, and operate by subtly changing the protein structure to reduce drug sensitivity without conferring an absolute resistance[6,45]. It will be important to determine whether the HIV-1 virus is equally adept at evading drugs that are directed against essential RNA targets, such as TAR and the RRE.

Drug resistance generally develops less rapidly for protease inhibitors than for RT inhibitors, but is still highly significant in clinical usage. This has limited the usefulness of protease-inhibitor monotherapy[14,46]. Distinct resistance mutants are obtained following treatment with transition-state analogues, C_2-symmetric analogues and the cyclic urea derivatives[8]. In some cases, the development of resistance to a drug has been rationalized in terms of molecular modelling. For example, Val84 is a target for mutation conferring low-level resistance to ABT-538 (Ref. 46). However, the broad cross-resistance shown by viruses obtained from patients treated with MK-639, to all members of a panel of six structurally diverse protease inhibitors, indicates that combination therapy with two or more protease inhibitors may be difficult to achieve[14].

In the search for improved treatment strategies, a number of combination clinical trials have been initiated in the hope that virus populations that are resistant to one drug might be sensitive to another. One approach has involved convergent therapy – two or more drugs are targeted to a single viral protein. This has only been possible because of the large number of drugs targeted to HIV-RT that are already licensed or in development. For example, in an initial 24-week trial, a combination of AZT and 3TC has shown substantially improved maintenance of viral-load reduction in HIV-infected patients compared with AZT alone[47]. Another attractive strategy to cope with resistance is to try to design inhibitors that are specifically targeted to commonly found resistant mutants. Undoubtedly, these discovery programmes will rely heavily on structural studies of drug-resistant enzymes[45].

Conclusions

There can be few tasks in biotechnology that are more challenging than designing antiviral drugs. All of the protease inhibitors that have entered into clinical trials are potent inhibitors of HIV-1 replication in cell culture, and exhibit remarkable selectivities for the viral enzyme. Unfortunately, early protease inhibitors tended to suffer from problems of short serum half-life, poor bioavailability and rapid clearance. As these pharmacokinetic problems have been addressed and solved, new difficulties have emerged from the resultant clinical experience, such as sequestration of the drug by serum proteins, drug resistance and uneven distribution throughout the body. Since these types of problems are unpredictable, it remains necessary to take into account the pharmacological parameters in any drug development programme at the earliest possible stage. Structure-based drug design clearly holds great promise for speeding-up the preclinical phases of drug development, but sadly it is not a panacea. Much hard work (and perhaps a little luck) will be needed before safe and effective anti-HIV drugs become a reality.

Acknowledgements

We should like to thank Jingshan Ren and David Stuart (University of Oxford) for providing the RT–drug complex structures shown in Fig. 2, Dale Kempf (Abbott, Pharmaceutical Products Division) for providing the protease–ABT-538 complex structures shown in Fig. 3 and for commenting on the HIV protease section, and Fareed Aboul-ela and Gabriele Varani (MRC Cambridge) for providing the TAR structure shown in Fig. 4.

References

1 Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. and Steitz, T. A. (1992) *Science* 256, 1783–1790
2 Jacobo-Molina, A. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 6320–6324
3 Rodgers, D. W. *et al.* (1995) *Proc. Natl Acad. Sci. USA* 92, 1222–1226
4 Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D. and Stuart, D. (1995) *Nat. Struct. Biol.* 2, 303–308
5 Ren, J. *et al.* (1995) *Nat. Struct. Biol.* 2, 293–302
6 Tantillo, C. *et al.* (1994) *J. Mol. Biol.* 243, 369–387
7 Spence, R. A., Kati, W. M., Anderson, K. S. and Johnson, K. A. (1995) *Science* 267, 968–993
8 Mellors, J. W., Larder, B. A. and Schinazi, R. F. (1995) *Int. Antiviral News* 3, 8–13
9 Wlodawer, A. *et al.* (1989) *Science* 245, 616–621
10 Miller, M. *et al.* (1989) *Science* 246, 1149–1152
11 Wlodawer, A. and Erikson, J. W. (1993) *Annu. Rev. Biochem.* 62, 543–585
12 West, M. L. and Fairlie, D. P. (1995) *Trends Pharmacol. Sci.* 16, 67–75
13 Roberts, N. A. *et al.* (1990) *Science* 248, 358–361
14 Condra, J. H. *et al.* (1995) *Nature* 374, 569–571
15 Kempf, D. J. *et al.* (1995) *Proc. Natl Acad. Sci. USA* 92, 2484–2488
16 Gehlhaar, D. K. *et al.* (1995) *Chem. Biol.* 2, 317–324
17 Kim, E. E. *et al.* (1995) *J. Am. Chem. Soc.* 117, 1181–1182
18 Lunney, E. A. *et al.* (1994) *J. Med. Chem.* 37, 2664–2677
19 Prasad, J. V. N. *et al.* (1995) *J. Med. Chem.* 38, 898–905
20 Thaisrivongs, S. *et al.* (1994) *J. Med. Chem.* 37, 3200–3204
21 Romines, K. R. *et al.* (1995) *J. Med. Chem.* 38, 1884–1891
22 Lam, P. Y. S. *et al.* (1994) *Science* 263, 380–384
23 Wong, Y. N. *et al.* (1994) *Biopharm. Drug Dispos.* 15, 535–544
24 Kuntz, I. (1992) *Science* 257, 1078–1082
25 Holloway, M. K. *et al.* (1995) *J. Med. Chem.* 38, 305–317
26 Hazuda, D. J., Hastings, D. J., Wolfe, A. L. and Emini, E. (1994) *Nucleic Acids Res.* 22, 1121–1122
27 Fesen, M. R., Kohn, K. W., Leteurtre, F. and Pommier, Y. (1993) *Proc. Natl Acad. Sci. USA* 90, 2399–2403
28 Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R.

and Davies, D. R. (1994) *Science* 266, 1981–1986
29 Gait, M. J. and Karn, J. (1993) *Trends Biochem. Sci.* 18, 255–259
30 Mann, D. A. *et al.* (1994) *J. Mol. Biol.* 241, 193–207
31 Graham, G. J. and Maio, J. J. (1990) *Proc. Natl Acad. Sci. USA* 87, 5817–5821
32 Sullenger, B. A., Gallardo, H. F., Ungers, G. E. and Gilboa, E. (1990) *Cell* 63, 601–608
33 Hsu, M-C. *et al.* (1991) *Science* 254, 1799–1802
34 Cupelli, L. A. and Hsu, M-C. (1995) *J. Virol.* 69, 2640–2643
35 Michne, W. F., Schroeder, J. D., Bailey, T. R., Young, D. C., Hughes, J. V. and Dutko, F. J. (1993) *J. Med. Chem.* 36, 2701–2702
36 Aboul-ela, F., Karn, J. and Varani, G. *J. Mol. Biol.* (in press)
37 Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D. and Williamson, J. R. (1992) *Science* 257, 76–80
38 Battiste, J. L., Tan, R., Frankel, A. D. and Williamson, J. R. (1994) *Biochemistry* 33, 2741–2747
39 Peterson, R. D., Bartel, D. P., Szostak, J. W., Horvath, S. J. and Feigon, J. (1994) *Biochemistry* 33, 5358–5366
40 Schröder, H. C., Ushijima, H., Bek, A., Merz, H., Pfeifer, K. and Müller, W. E. G. (1993) *Antiviral Chem. Chemother.* 4, 103–111
41 Zapp, M. L., Stern, S. and Green, M. R. (1993) *Cell* 74, 969–978
42 Ciccarelli, R. B. *et al.* (1994) *Antiviral Chem. Chemother.* 5, 169–175
43 Coffin, J. M. (1995) *Science* 267, 483–489
44 Ho, D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. and Markowitz, M. (1995) *Nature* 373, 123–126
45 Erickson, J. W. (1995) *Nat. Struct. Biol.* 2, 523–529
46 Markowitz, M. *et al.* (1995) *J. Virol.* 69, 701–706
47 Katlama, C. and The European Lamivudine HIV Working Group (1994) *Abstracts of the Int. Congress on Drug Therapy in HIV Infection*, Glasgow, UK

Recent developments in retro peptides and proteins – an ongoing topochemical exploration

Michael Chorev and Murray Goodman

Main-chain peptidomimetics based on peptide-bond reversal and inversion of chirality represent important structural alterations for peptides and proteins, and are highly significant for biotechnology; these modifications have been widely applied: the D-HIV-protease dimer cleaves only all-D substrate; an all-D-hexapeptide opioid is able to produce analgesia following intraperitoneal administration. Antigenicity and immunogenicity can be achieved by metabolically stable antigens such as all-D- and retro-inverso-isomers of natural antigenic peptides. Isomers, including the retro- and retro-inverso- forms, of hybrid peptides derived from cercropin A and melittin, maintain antimicrobial activity. Therefore, an insight is provided into structure–activity relationships and the rational design of biologically important isomeric peptides.

The non-palindromic nature of peptide and protein structures remains the basis for the synthesis of mimetics and analogs of natural bioactive molecules. The complicated and subtle interplay between the backbone and side chains provides a fertile ground for studies attempting to relate structure to physicochemical properties, conformational preferences and biological functions. It is well established that molecular recognition, the epitome of specific and efficacious biochemical processes, involves interaction between complementary molecular surfaces. Therefore, the identification and characterization of relevant molecular topology is central to 'rational drug design', the ultimate goal of which is to mimic this putative 'bioactive topology'. This is accomplished by developing partially peptidic, pseudopeptidic and non-peptidic molecular entities that are recognized and are able to induce the characteristic behaviors of the parent peptides[1–4]. The combination of 'sense of direction'

M. Chorev is at the Department of Pharmaceutical Chemistry, Faculty of Medicine, The Hebrew University, Jerusalem, Israel. M. Goodman is at the Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093-0343, USA.